

Technical Report

Case Study: Video Games Sales

Problem Statement

The video game industry is a rapidly growing and highly competitive market. Understanding the factors that influence video game sales is crucial for developers, publishers, and retailers to make informed decisions about game development, pricing, and marketing strategies. The objective of this analysis is to investigate the relationship between various factors, such as genre, platform and their impact on video game sales.

Dataset

Here is the sample from the dataset:-

Platform	Genre	Global_Sales (in millions)
X360	Racing	1.52
PS3	Racing	6.66
PC	Shooter	2.06
Wii	Shooter	5.46
PS3	Shooter	10.34
PC	Action	0.90
PC	Shooter	0.09
PC	Racing	0.19
Wii	Action	8.41
X360	Shooter	37.50

Data consists of following values:

- Platforms : PC, PS3, Wii, X360
- Genre : Action, Racing, Shooter

The data consists of samples for a given *Genre* and given *Platform* what is the *Global_Sales*. Each sample represents sales in a particular year.

Hypothesis

HO1 : Global Sales is not affected by different Genres

HA1 : Global Sales is affected by different Genres

HO2 : Global Sales is not affected by different Platforms

HA2 : Global Sales is affected by different Platforms

HO2 : Global Sales is not affected by interaction of Genre and Platform

HA2 : Global Sales is affected by interaction of Genre and Platform

ANOVA Testing

Using Python and “*statsmodel*” library we were able to perform ANOVA 2-way testing on the dataset.

The code is available on the github (mentioned in references).

	df	Sum of Squares	Mean of Squares	F	P
Genre	2	2319.852	1159.926	9.74487	$1.28 * 10^{-4}$
Platform	3	5969.652	1989.884	16.7175	$5.43 * 10^{-9}$
Platform+ Genre	6	2090.561	348.426	2.92723	$1.10 * 10^{-2}$
Residual	108	12855.17	119.0293		

Results and Interpretations

- We see that ***p-value*** for *Genre* is < 0.05 , hence we can say that there is significant difference between different *Genres* which affect the *Global_Sales*.
- This indicates that Genre like Shooter, Action, Adventure have significant impact on the Global Sales compared to Genres like Simulation, Puzzle, etc.
- Hence, we reject **H01** and accept **HA1**.
- We see that the ***p-value*** for *Platform* is also < 0.05 , hence we say that there is a significant difference between different Platforms which affect the *Global_Sales*.
- We can interpret it as there are more sales in Consoles (like XBOX 360 and PS3) compared to PC, Wii.
- Hence, we reject **H02** and accept **HA1**.
- We see that the ***p-value*** for *Platform + Genre* is > 0.05 , hence we say that there is no significant difference between *Platform + Genre* which affects the *Global Sales*.
- It indicates that the sales made by games for a platform and a genre does not vary significantly.
- Hence, we accept **H03**.

References

- Github Repository for code :
<https://github.com/QuickHawk/anova-vg>
- Video Game Sales Dataset :
<https://www.kaggle.com/datasets/gregorut/videogamesales>